

doi:10.11835/j.issn.1000-582X.2020.309

办公建筑运行能耗的混沌时间序列复合预测

于军琪,段佳音,赵安军,井文强,王佳丽

(西安建筑科技大学 建筑设备科学与工程学院,西安 710055)

摘要:针对办公建筑已有的能耗预测方法中未能考虑到能耗数据的混沌变化特性,提出了一种基于混沌时间序列的办公建筑运行能耗预测方法。对研究对象的时间序列进行相空间重构,判断其具备混沌特性,建立混沌理论和支持向量回归的组合模型进行训练,采用 Markov 链消除组合模型由于参数传递产生的累积误差,得到最终预测结果。为了验证算法的有效性,以西安某办公建筑的能耗监测数据为例进行实例分析,并与非线性自回归神经网络、支持向量回归等其他预测方法进行对比。实验结果表明,经过 Markov 修正后的混沌时间序列组合模型预测精度显著提高,预测效果优于其他方法,且更符合办公建筑能耗的变化规律,为节能优化提供有效的数据支撑。

关键词:办公建筑能耗;混沌时间序列;预测算法;马尔科夫链

中图分类号:TU831

文献标志码:A

文章编号:1000-582X(2021)09-077-11

Chaotic time series composite prediction of office building energy consumption

YU Junqi, DUAN Jiayin, ZHAO Anjun, JING Wenqiang, WANG Jiali

(School of Building Services Science and Engineering, Xi'an University of Architecture and Technology, Xi'an 710055, P. R. China)

Abstract: The existing energy consumption prediction methods for office buildings fail to take into account the chaotic change characteristics of energy consumption data. In this paper, a method of energy consumption prediction for office buildings based on chaotic time series was proposed. The method first reconstructed the phase space of the time series of the research object, and judged that whether it had chaotic characteristics. Then the combination model of chaos theory was established and applied in vector regression for training. Finally, Markov chain was used to eliminate the cumulative errors caused by parameter transfer of the combination model, and the final prediction result was obtained. In order to verify the effectiveness of the algorithm, the energy consumption monitoring data of an office building in Xi'an was taken as an example for analysis. The proposed method was compared with other prediction methods, such as nonlinear autoregressive neural network and support vector regression. The experimental results

收稿日期:2020-01-05 网络出版日期:2020-03-06

基金项目:安徽建筑大学智能建筑与建筑节能安徽省重点实验室开放课题资助项目(Z20190383);碑林区应用技术研发类资助项目(GX1903)。

Supported by the Open Foundation of the Key Lab (Center) of Intelligent Building and Building Energy Conservation, Anhui Jianzhu University (Z20190383), and Beilin Area Applied Technology Research and Development Project(GX1903).

作者简介:于军琪(1969—),男,博士生导师,主要从事智能建筑研究,(E-mail)928264474@qq.com。

通讯作者:赵安军,男,副教授,(E-mail)zhao_anjun@163.com。

show that the prediction accuracy of the chaotic time series combination model modified by Markov was significantly improved, the prediction result was better than those of other models and more consistent with the change law of energy consumption of office buildings, providing effective data support for energy conservation optimization.

Keywords: energy consumption of office buildings; chaotic time series; prediction algorithm; Markov chain

办公建筑在全生命周期中普遍存在能耗大、能效低等问题,导致严重的能源浪费。其节能潜力巨大,降耗幅度可以达到30%~50%^[1]。其中,办公建筑运行能耗占比最大。因此,研究办公建筑的运行能耗情况具有重要意义,对其进行实时准确地预测可以为优化运行效率提供数据决策,从而实现节能目标。

建筑能耗的预测方法主要分为两大类:正演模型和数据驱动模型。笔者根据获得的办公建筑历史能耗数据特点,利用数据驱动模型中的时间序列法进行能耗短期预测。机器学习预测方法是最常用的短期能耗预测方法,已经广泛应用于风速预测^[2]、电力需求预测^[3]、建筑能耗预测^[4]、建筑冷负荷预测^[5]等能源领域。目前国内外建筑能耗时间序列预测方法应用较为广泛的有:人工神经网络^[6](ANN, artificial neural network)、差分整合移动平均自回归^[7](ARIMA, autoregressive integrated moving average)、支持向量回归^[8](SVR, support vector regression)、多元线性回归^[9](MLR, multiple linear regression)等。办公建筑除了室外气象参数和围护结构等本体能耗外,还有机电设备、节假日、人员占用率等复杂的运行能耗,造成其能耗时间序列具有非线性特征。为了深入挖掘其时间序列内部的非线性规律,提高预测精度,研究者们将不同的算法与人工神经网络和支持向量机结合,建立了多种办公建筑混合预测模型。文献[9]根据天气、人员等多种因素利用MLR方法对办公建筑的逐日冷负荷进行了回归预测,与实际负荷相比平均绝对百分误差小于8%。文献[10]利用反馈神经网络(BPNN, back propagation neural network),径向基函数神经网络(RBFNN, radial basis function neural network),广义回归神经网络(GRNN, generalized regression neural network),支持向量机(SVM, support vector machine)方法分别对办公建筑的逐时能耗进行分析,以中国广州某办公楼为实例,取得了较好的预测效果。文献[11]建立了小波-支持向量机(wavelet-SVM),小波-偏最小二乘-支持向量机(Wavelet-PLS-SVM, wavelet-partial least squares-support vector machine)模型对办公建筑的逐时能耗进行了分析,最终得出1 h, 2 h, 3 h和24 h的预测结果。在其他应用领域^[12]将平均速度、平均占有率和平均交通流时间序列利用贝叶斯理论集成到一个时间序列中,将混沌理论和SVR相结合用于交通流的时间序列预测,达到了较高的预测精度。上述研究都是基于多变量输入的,相较于单变量时间序列需要考虑的因素复杂,文献[13]使用历史时刻的办公建筑冷负荷根据贝叶斯理论映射到同一时间维度作为机器学习预测模型的单变量输入,利用Chaos-SVR和WD-SVR(wavelet decomposition-support vector regression)分别对建筑冷负荷时间序列进行预测,结果表明,Chaos-SVR效果更优。因此,选择恰当的预测模型不仅可以避免单变量输入的局限性,而且会取得较好的预测效果,但由于组合模型在迭代过程中存在误差累积的问题,导致不能很好地满足实际优化运营的需求。

为了解决以上问题,笔者探究了混沌现象背后隐藏的决定性规律使得将混沌理论应用于办公建筑运行能耗预测领域的可能性,展开了以下研究工作:

- 1) 分析办公建筑运行能耗的特点及构成,进行数据预处理和时间序列的混沌特性判别。
- 2) 将混沌理论(Chaos theory)和支持向量理论结合,构建Chaos-SVR办公建筑能耗预测模型。
- 3) 利用Markov链修正Chaos-SVR组合模型中由于参数传递存在的累积误差。
- 4) 以西安某办公建筑监测数据为例进行实例分析,并与其他预测方法进行比较,验证了模型的有效性。

文中主要研究办公建筑运行能耗的混沌特性,其他大型公共建筑比如商业建筑、医疗建筑等能耗是否具有混沌特性,模型的可推广性将是下一步的研究重点。

1 能耗数据特征分析

办公建筑能耗混沌时间序列预测研究的前提是确定该时间序列具有混沌特征^[14]。文中通过西安市建

筑能耗监测平台,采集某办公建筑 2019 年 8 月 1 日 0 点到 9 月 30 日 23 点各项分类能耗数据,如表 1 所示。

表 1 西安某办公建筑能耗量统计

Table 1 Energy consumption statistics of an office building in Xi'an

时刻/ (年/月/日/时)	水耗量/ t	电耗量/ (kW·h)	燃气量/ m ³	集中供冷/热量/ GJ	其他能源使用量/ (kW·h)	综合能耗/ kgce
2019/8/1/00	8 001.23	659.23	30.56	21.21	2.34	689.40
2019/8/1/01	8 003.95	500.35	42.31	8.35	1.01	686.31
...
2019/9/30/22	8 007.13	1 111.24	50.26	7.68	0.95	686.64
2019/9/30/23	8 003.32	766.76	35.85	5.65	0.34	686.19

从表 1 可以看出,办公建筑运行能耗包括水耗量、电耗量、燃气量、集中供冷/热量和其他能源使用量。为了对办公建筑运行能耗时间序列进行混沌特性分析,由于其组成结构复杂,需要对多项数据统一量纲。因此,将所有运行能耗的量纲转化为千克标准煤(kgce)的形式,从而得到该办公建筑运行能耗的逐时序列。然后对该逐时序列数据进行处理,利用最大 Lyapunov 指数法来判断其是否具有混沌特性。最大 Lyapunov 指数法是基于相空间重构的,因此对时间序列进行相空间重构是研究的前提。

1.1 相空间重构

相空间是一种工具,用于描述由确定自由度的确定性混沌系统生成的单变量或多变量时间序列重构的动态系统^[15]。相空间重构的主要原理是在拓扑等价意义下从高维空间中恢复混沌吸引子的动力学特性^[15],对于总能耗的逐时时间序列 $x_1, x_2, x_3, \dots, x_{n-1}, x_n$,适当选取嵌入维数 m 和延迟时间 τ ,得到重构相空间

$$Y_m(i) = [x(i), x(i + \tau), x(i + 2\tau), \dots, x(i + (m - 1)\tau)], \tag{1}$$

式中, $i = 1, 2, \dots, n; N = n - (m - 1)\tau$ 。 N 表示由时间序列构成 m 维后,由 n 个相点变成了 N 个相点。

嵌入维数和延迟时间的合理选取决定了办公建筑能耗时间序列相空间重构的质量,同时也保证了在实际预测中有足够高的预测精度。研究表明^[16],将 τ 和 m 联合起来的嵌入窗宽的确定也会影响重构相空间的质量。C-C^[17]法通过序列的关联积分来构成统计量,通过统计量和延迟时间的关系图同时计算出延迟时间 τ 和最佳窗宽^[16] τ_w ,再根据 $\tau_w = (m - 1)\tau$ 求出嵌入维数。该方法既能有效地减少互信息量法的计算量,又能保持时间序列的非线性特征。故本文采用 C-C 法求取延迟时间和嵌入维数,具体步骤如下:

1)将办公建筑能耗时间序列 $\{x_i\}, i = 1, 2, \dots, n$ 分成 t 个不相交的时间序列,当 $t = 1$ 时为时间序列本身,当 t 为一般的自然数时,时间序列长度 $l = \frac{N}{t}$ 。构造办公建筑能耗时间序列的关联积分函数为

$$C(m, n, r, t) = \frac{2}{N(N-1)} \sum_{1 \leq i < j \leq N} \theta(r - d_{ij}), r > 0, \tag{2}$$

式中, $d_{ij} = \|X_i - X_j\|$, X_i 为相空间的点, $\theta(\cdot)$ 为阶跃函数:

$$\theta(\cdot) = \begin{cases} 0, & x < 0, \\ 1, & x \geq 0. \end{cases} \tag{3}$$

2)统计量代表了非线性时间序列的相关性。通过时间序列的关联积分来构成每个不相交的时间子序列的检验统计量:

$$S(m, n, r, t) = \frac{1}{t} \sum_{s=1}^t [C_s(m, N/t, r, t) - C_s^m(1, N/t, r, t)], \tag{4}$$

令 $n \rightarrow \infty$, 则:

$$S(m, r, t) = \frac{1}{t} \sum_{s=1}^t [C_s(m, r, t) - C_s^m(1, r, t)], (m = 2, 3, \dots). \tag{5}$$

3)如果时间序列是独立同分布的,对于固定的 m 和 t ,当 $N \rightarrow \infty$ 时, r 取任意值时, $S(m, r, t)$ 恒等于零。

但实际的时间序列是有限的,因此一般 $S(m, r, t)$ 不等于零。那么,局部最大时间间隔可以取 $S(m, r, t)$ 的零点或对所有的半径 r 相互差别最小的时间点,这意味着这些点几乎是均匀分布的。选择对应值最大和最小的 2 个半径 r , 定义关于 r 的最大偏差:

$$\Delta S(m, t) = \max\{S(m, r_j, t)\} - \min\{S(m, r_j, t)\}。 \quad (6)$$

局部最大时间 t 应该是 $S(m, r, t)$ 的零点和 $\Delta S(m, t)$ 的最小值。但是 $S(m, r, t)$ 的零点对所有 m, r 应几乎相等; $\Delta S(m, t)$ 的最小值对所有 m 应几乎相等。延迟时间 τ 对应着这些局部最大时间 t 中的第一个。

4) 通过统计量和延迟时间的关系图来确定 τ 和嵌入窗宽 τ_w , 从而确定出嵌入维数^[16]。

几种重要渐进分布的数学统计结果表明: 当 $2 \leq m \leq 5, \sigma/2 \leq r \leq 2\sigma, N \geq 500$ 时 (σ 指时间序列的均方差或标准差), 渐近分布可以通过有限序列很好的近似, $S(m, n, r, 1)$ 能代表序列的相关性。具体计算时, 根据办公建筑能耗时间序列特性, 取 $m = 2, 3, 4, 5, r_i = i\sigma/2, i = 1, 2, 3, 4$, 相点数 $n = 3\ 000$, 计算下列 3 个检验统计量^[17]:

$$\bar{S}(t) = \frac{1}{16} \sum_{m=2}^5 \sum_{j=1}^4 S(m, r_j, t), \quad (7)$$

$$\Delta \bar{S}(t) = \frac{1}{4} \sum_{m=2}^5 \Delta S(m, t), \quad (8)$$

$$S_{\text{cor}}(t) = \Delta \bar{S}(t) + |\bar{S}(t)|。 \quad (9)$$

根据上述计算结果在 MATLAB 中绘制图像, 得到 C-C 法重构办公建筑能耗时间序列统计量曲线如图 1 所示。

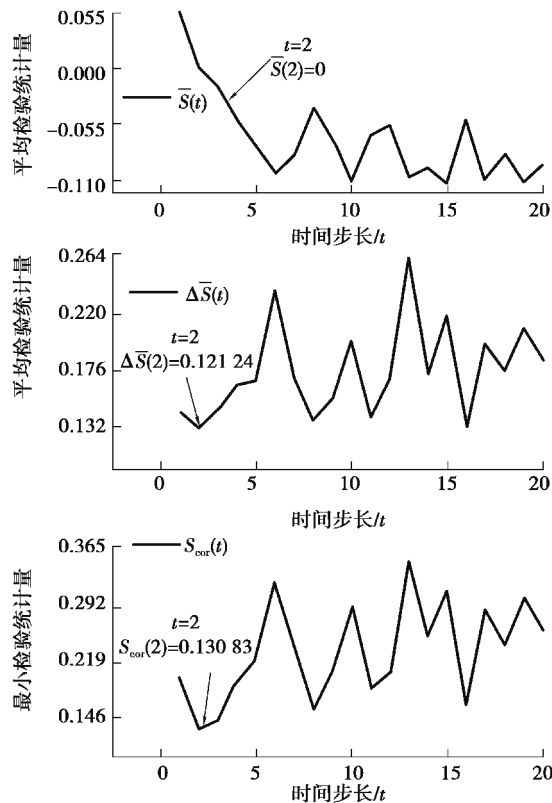


图 1 C-C 法重构办公建筑总能耗时间序列统计量曲线图

Fig. 1 Time series statistics curve of total energy consumption reconstructed by C-C method

分析得到, $\bar{S}(t)$ 第一个零点或 $\Delta \bar{S}(t)$ 第一次取得的极小值, 时间序列独立的第一个局部最大值的对应的延迟时间即为最佳延迟时间 $\tau = 2$ 。 $S_{\text{cor}}(t)$ 取得最小值, 时间序列独立的第一个整体最大时间窗口 t 为最佳窗宽 $\tau_w = 2$, 根据嵌入时间窗宽公式 $\tau_w = (m - 1)\tau$ 可确定嵌入维数 $m = 2$ 。

因此,该办公建筑总能耗时间序列重构的相空间为 $Y_2(i)=[x(i),x(i+2)]$ 。

1.2 混沌特性判别

一个系统的 Lyapunov 指数谱中有一个或一个以上的正值就可以肯定其存在混沌特性^[18]。在相空间重构后,相空间中的两条相邻轨线将随时间逐渐发散或聚合,而 Lyapunov 指数是其轨线的收敛或发散率,正的最大 Lyapunov 指数反映时间序列具有混沌特征,负的最大 Lyapunov 指数表明时间序列具有随机性或周期性。最大 Lyapunov 指数越大,时间序列的非线性越强,对初值越敏感^[18]。

小数据量法^[19]可以充分利用所有能够利用的数据,对小数据组来说比较可靠,计算量不大,相对容易操作,计算结果精度高。因此采用小数据量方法计算最大 Lyapunov 指数。小数据量法求取最大 Lyapunov 指数的步骤如下:

1)平均周期是从相空间的意义上度量一个吸引子全局属性的最小单元,利用快速傅里叶变换得到频谱信息,计算出能耗的平均频率根据其倒数估算出时间序列的平均周期 τ_p 。

经过傅里叶变换后,平均周期可表示为

$$\tau_p = \sum_{i=1}^L P_i \left(\sum_{i=1}^L f_i P_i \right) - 1, \quad (10)$$

式中: $f_i(i=1,2,\dots,L)$ 是 0 Hz 和奈奎斯特频率之间均匀分布的频率点; P_i 是在频率点 f_i 处的幅值。用C-C法计算出延迟时间 τ 和嵌入维数 m 。对于办公建筑能耗时间序列 $\{x_i\}$,重构后的相空间为式(1)。

2)寻找重构相空间中每个点 $Y_m(t)$ 的最近邻点 $Y_m(\hat{t})$,并限制短暂分离:

$$L_t(0) = \min_{\hat{t}} \| Y_m(t) - Y_m(\hat{t}) \|, |t - \hat{t}| > \tau_p, \quad (11)$$

式中: $\hat{t}=1,2,\dots,N,t \neq \hat{t}$; τ_p 为办公建筑能耗时间序列的平均周期; $L_t(0)$ 是初始时刻一对最近邻点之间的距离。

3)根据相空间中每个点的最邻近点的平均发散速率估计最大 Lyapunov 指数为

$$\lambda(i) = \frac{1}{i\Delta t} \frac{1}{(N-i)} \sum_{j=1}^{N-i} \ln \frac{L_t(i)}{L_t(0)}, \quad (12)$$

式中: Δt 为样本周期。对相空间中的每个点 $Y_m(t)$,计算出该邻域点经过 i 个离散事件步长后的距离

$$L_t(i) = \| Y_m(t+i) - Y_m(\hat{t}+i) \|, i=1,2,\dots,\min(N-t,N-\hat{t}), \quad (13)$$

估计得到:

$$L_t(i) = C_j e^{\lambda(i\Delta t)}, C_j = L_t(0). \quad (14)$$

最大 Lyapunov 指数的几何意义是量化初始闭轨道的指数发散和估计系统的总体混沌水平的量^[11],对式(14)两边取对数得到

$$\ln L_t(i) = \ln C_j + \lambda(i\Delta t) (i=1,2,\dots,N). \quad (15)$$

4)对每个 i 求出所有 t 的 $\ln L_t(i)$ 平均值 $y(i)$,即:

$$y(i) = \frac{1}{q\Delta t} \sum_{j=1}^q \ln L_t(i), \quad (16)$$

式中: q 是非零 $L_t(i)$ 的数目。最大 Lyapunov 指数相当于式(15)直线的斜率,可以通过最小二乘法逼近这条直线得到。

根据相空间重构 C-C 法求出的最佳延迟时间 $\tau=2$ 及嵌入维数 $m=2$,利用小数据量法计算得出如图 2 所示办公建筑能耗时间序列的最小二乘法拟合直线。

根据直线斜率可以得到最大 Lyapunov 指数值为 0.007 7,此为正值,由此可得出办公建筑总能耗量时间序列具有混沌特性。

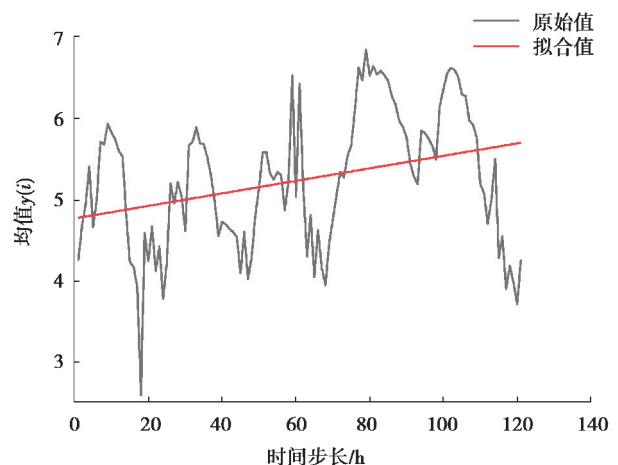


图 2 最小二乘法拟合直线

Fig. 2 Line fitting by least square method

2 Chaos-SVR 组合模型预测

办公建筑的 Chaos-SVR 的预测方法主要运用混沌理论,构造非线性映射利用相空间重构技术将原始时间序列数据所隐含的信息进行还原重构,并将其作为 SVR 模型的输入,进行非线性训练和预测。基于办公建筑能耗的 Chaos-SVR 时间序列模型建立及预测过程如下:

1) 构建预测模型结构,主要包括 3 部分:输入向量、支持向量的非线性变化和输出。西安某办公建筑运行能耗的时间序列进行重构后的相空间为 $Y_2(i) = [x(i), x(i+2)]$,相点数为 $N = n - 2$,每个相点包含了每个序列的主要特征,能够逼近建筑能耗的真实状态。向量 $x_i = (x_i, l)$ 是模型的输入向量, $\beta = (\beta_1, \beta_2, \dots, \beta_l)$ 是输出的权重向量, $y_i = f(x_i)$ 是模型的输出。Chaos-SVR 模型结构如图 3 所示。

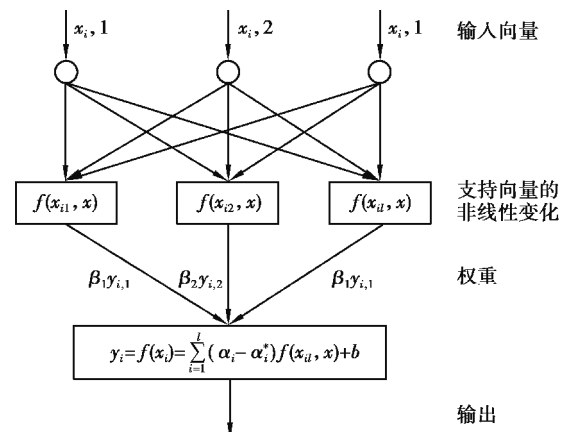


图 3 Chaos-SVR 模型结构图

Fig. 3 Chaos-SVR model structure diagram

2) 利用支持向量机中的 ϵ 不敏感损失函数,得到 ϵ -SVR 模型,采用高斯径向基函数 $f(x_i, x) = \exp(-\gamma \|x_i - x\|^2)$ 作为 ϵ -SVR 的核。 ϵ -SVR 估计函数的形式为: $y_i = f(x_i) = \sum_{i=1}^l (\alpha_i - \alpha_i^*) f(x_i, x) + b$, 其中 α_i, α_i^* 是 lagrange 乘子。对于 ϵ -SVR, 一个重要的问题是模型参数的选择,采用遗传算法 (genetic algorithm, GA) 来优化惩罚因子 C 和核函数参数 γ 。

3) 进行网络循环学习,修正 SVR 的输出及权值,直到误差控制在允许范围或迭代次数达到上限,训练结束。

4) 取样本前 n 个数据之后连续的 Q 个数据点作为预测数据。对模型进行预测,网络的输出值即为预测值,将预测值进行相应的反归一化处理,得到预测的实际值。

为了验证模型的有效性,选取 2019 年 9 月 4 日~9 月 24 日共计 20 天 480 组数据作为训练数据,以 9 月 25 日和 26 日 2 天共计 48 组数据作为预测数据来进行模型验证。并与目前在时间序列预测领域应用广泛的 2 种机器学习方法——非线性自回归神经网络 (nonlinear auto regressive neural network) 和支持向量回归 (SVR) 方法的预测值和实际值进行比较,利用均方根误差 RMSE (root mean squared error) 平均绝对百分比误差 MAPE (mean absolute percent error) 作为评价指标。预测对比结果如图 4 所示。

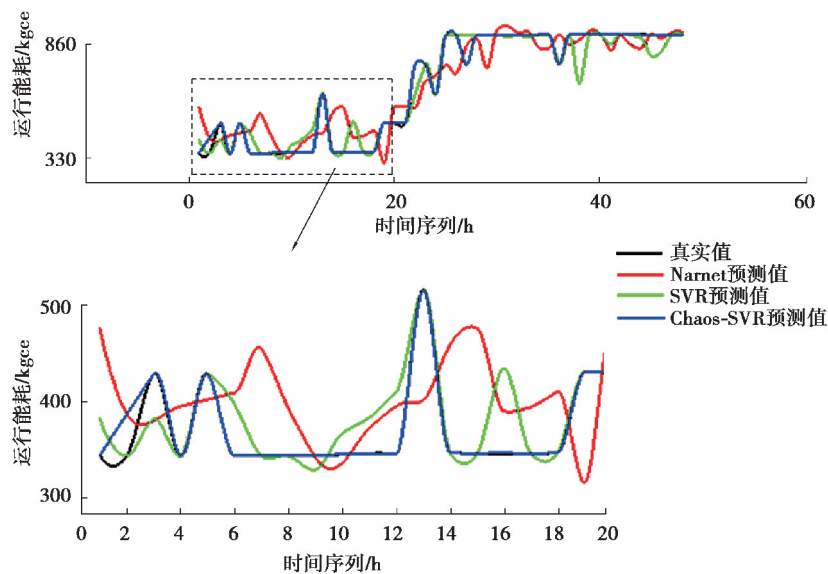


图 4 预测结果对比图

Fig. 4 Comparison of prediction results

预测结果误差对比如图 5 所示。

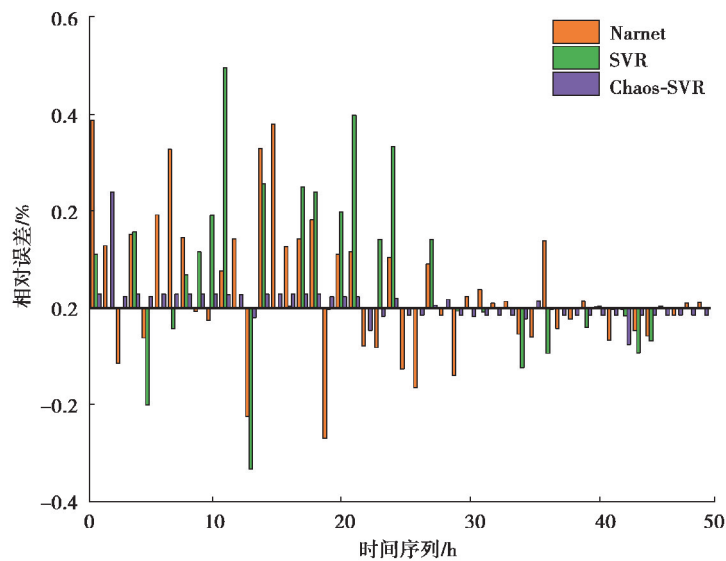


图 5 预测结果误差对比图

Fig. 5 Error comparison diagram of prediction results

预测结果误差如表 2 所示。

表 2 预测模型误差值

Table 2 Error values of prediction model

采用方法	RMSE	MAPE
Narnet	38.321 2	10.569 5
SVR	23.928 6	3.415 8
Chaos-SVR	6.326 1	0.287 5

从图 4 中可以看出, Narnet 预测曲线波动最大, SVR 预测曲线较为平缓, Chaos-SVR 预测曲线距离真实值曲线最接近, 部分值几乎完全拟合。由图 5 得, Narnet 预测方法误差最大, SVR 预测方法误差较大, Chaos-SVR 预测方法误差在 0 值附近小范围浮动。由表 2 可以看出, Chaos-SVR 预测模型的 RMSE 为 6.326 1, MAPE 为 0.287 5, 相较于其他 2 种方法均最小。实验是在同样的环境和同一组数据的基础上进行的, Chaos-SVR 预测方法相较于 Narnet 和 SVR 方法, 预测效果最好, 预测精度明显提高。但 Chaos-SVR 组合模型在预测的过程中, 还有部分值存在较大的相对误差, 为了更加精确地进行办公建筑运行能耗预测, 需要对这部分的误差进行修正。导致这种误差存在的原因可能是由于组合模型在迭代过程中, 由于参数传递而出现的过程误差积累, 这样的误差积累会造成预测效果存在一定的缺陷。为了进一步提高预测精度, 文中利用 Markov 链进行修正。

3 Markov 链修正误差

Markov 链是一种具有无后效性的随机过程^[20], Markov 链适用于对存在波动性的预测问题进行修正描述^[21]。它可以根据某一已知条件的时刻条件下推求出下一时刻的概率分布, 得到下一时刻的状态而与其他时刻无关。因此, 在传统 Chaos-SVR 预测模型基础之上利用 Markov 链对累积误差进行修正, 从而实现办公建筑能耗时间序列的精确预测。

1) 误差状态区间^[22]划分。

误差状态区间表示未来时刻预测值的误差状态, 用 E_1, E_2, \dots, E_m 分别来表示, 并且状态转移只在 t_1, t_2, \dots, t_m 等可数时刻发生。定义误差均值 \bar{X} 和标准差 S 分别为

$$\bar{X} = \frac{1}{n} \sum_{k=1}^n \bar{A}(t), \quad (17)$$

$$S = \sqrt{\frac{1}{n-1} \sum_{k=1}^n (\bar{A}(t) - \bar{X})^2}, \quad (18)$$

式中: $\bar{A}(t)$ 为 Chaos-SVR 预测模型所求出的预测值矩阵, $A(t)$ 为真实值矩阵, 根据中心极限原理^[15], 将误差状态区间划分为 5 组, 即 $(-\infty, \bar{X}-S)$, $[\bar{X}-S, \bar{X}-0.5S)$, $[\bar{X}-0.5S, \bar{X}+0.5S)$, $[\bar{X}+0.5S, \bar{X}+S)$, $[\bar{X}+S, \infty)$ 。

2) 状态转移概率矩阵。

Markov 链由状态 E_i 经过 k 步转移到状态 E_j 的转移概率用 $P_{ij}^{(k)}$ 表示, 其含义是指未来时刻会处于哪一个误差状态区间, 为

$$P_{ij}^{(k)} = \frac{n_{ij}^{(k)}}{N_i}, \quad (19)$$

式中: N_i 表示状态 E_i 出现的总次数; $n_{ij}^{(k)}$ 表示状态 E_i 误差经过 k 步转移到状态 E_j 的次数; n 为划分的状态区间数目。那么一步状态转移概率矩阵为

$$P^{(1)} = \begin{bmatrix} P_{11} & P_{12} & \cdots & P_{1m} \\ P_{21} & P_{22} & \cdots & P_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ P_{m1} & P_{m2} & \cdots & P_{mm} \end{bmatrix}, \quad (20)$$

使用 C-K 方程(切普曼-柯尔莫哥洛夫方程)计算 k 步的转移概率矩阵为

$$P(k) = (P^{(1)})^k. \quad (21)$$

3) 预测模型。

设初始向量为 $P(0)$, 经 k 步转移后的状态向量为

$$P(k) = P(0) \times P(k) = P(0) \times (P^{(1)})^k, \quad (22)$$

据此可以判断 k 步所处的状态空间。

4) 根据步骤 2) 所得概率转移矩阵和步骤 3) 所得状态向量, 得出误差修正后的值。

以该办公建筑 9 月 25 日的预测数据为例进行 Markov 链参数的设置。首先通过预测能耗值与实际能耗值的误差均值 $\bar{X}=0.008$ 和标准差 $S=0.088$ 确定误差状态空间为: $E_1 = (-\infty, -0.08)$, $E_2 = [-0.08, -0.036)$, $E_3 = [-0.036, 0.052)$, $E_4 = [0.052, 0.096)$, $E_5 = [0.096, +\infty)$ 。

根据划分的状态区间对 Chaos-SVR 模型预测值的误差值进行状态划分, 划分结果如表 3 所示。

表 3 Chaos-SVR 模型预测误差分类结果

Table 3 Prediction error classification results of Chaos-SVR model

时间	误差	状态	时间	误差	状态	时间	误差	状态
0	0.011 1	E_3	8	0.016 1	E_3	16	-0.001 3	E_3
1	0.496 0	E_5	9	0.091 5	E_4	17	0.004 3	E_3
2	-0.000 3	E_3	10	0.000 4	E_3	18	0.023 9	E_3
3	0.057 6	E_3	11	-0.031 4	E_1	19	-0.003 1	E_3
4	-0.099 7	E_1	12	0.025 6	E_3	20	0.078 0	E_4
5	-0.000 4	E_3	13	-0.001 3	E_3	21	0.039 8	E_3
6	-0.042 1	E_2	14	0.004 3	E_3	22	0.142 0	E_5
7	0.069 0	E_4	15	0.025 0	E_3	23	0.001 5	E_3

根据上述误差状态区间的确定可以得到 Chaos-SVR 预测结果的 Markov 链的状态转移情况, 进一步计算出状态转移概率矩阵 P 为

$$P = \begin{bmatrix} 0 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{1} & 0 \\ \frac{2}{16} & \frac{1}{16} & \frac{8}{16} & \frac{1}{16} & \frac{2}{16} \\ 0 & 0 & \frac{3}{3} & 0 & 0 \\ 0 & 0 & \frac{2}{2} & 0 & 0 \end{bmatrix} \quad (23)$$

根据 Markov 链预测模型得到办公建筑 9 月 25 日的能耗预测状态向量,针对 Chao-SVR 预测模型存在的累积误差进行修正,得到该办公建筑能耗的最终预测值。为了验证基于 Markov 链修正后的 Chaos-SVR 预测效果,对修正前后的预测结果进行比较。对比结果如图 6 所示。

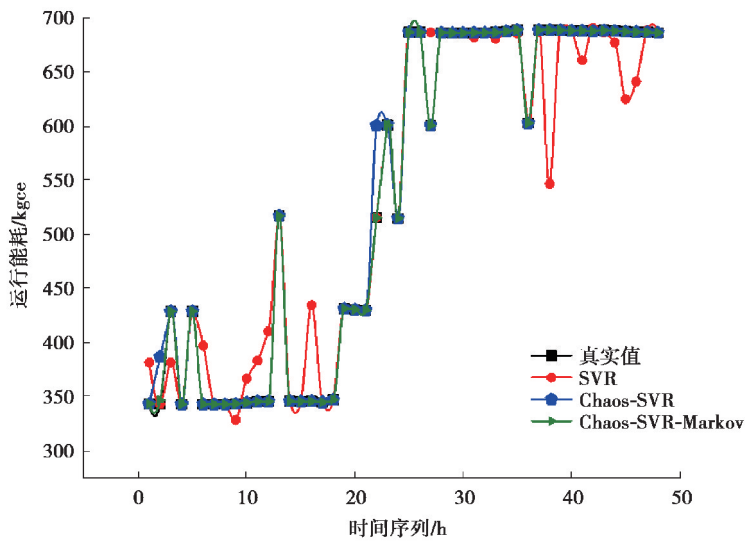


图 6 Markov 修正前后预测结果对比图

Fig. 6 Comparison of Markov's prediction results before and after revision

Markov 链修正前后预测结果误差对比如图 7 所示。

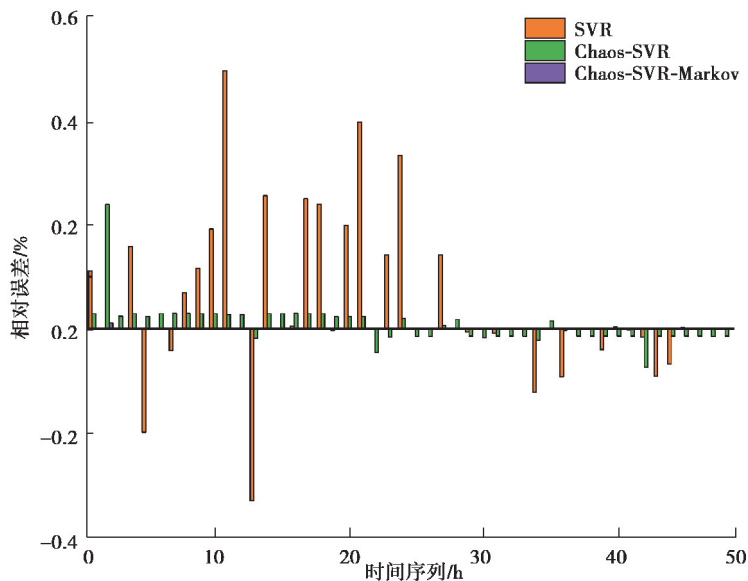


图 7 Markov 链修正前后预测结果误差对比图

Fig. 7 Error comparison diagram of prediction results before and after Markov chain correction

Markov 链修正前后的误差计算结果如表 4 所示。

表 4 Markov 链修正前后误差值

采用方法	RMSE	MAPE
Chaos-SVR	6.326 1	0.287 5
Chaos-SVR-Markov	0.565 5	0.044 1

由图 6 得,经 Markov 修正后的模型预测曲线和真实值曲线拟合度很高,直观地可以看到主要对 2 个时间点的运行能耗值进行了修正。从图 7 预测结果误差对比中可以得到证实,经过 Markov 链修正之后得误差非常小,不仅对部分累积误差进行修正接近于 0 值,其他误差也进行了相应修正,总体预测精度显著提高。由表 4 可以看到,经 Markov 修正之后的 RMSE 和 MAPE 均降低,这说明经过 Markov 链对组合模型中参数传递存在的累积误差进行了有效改进,它的修正过程适用于办公建筑运行能耗这一非线性系统。Markov 链的无效性特征在组合模型的误差修正中得到了充分体现,只对最后的预测结果做处理而不影响传统预测模型的运算过程,更符合实际办公建筑运行能耗的变化规律。因此,与 Chaos-SVR 时间序列预测模型相比,基于 Markov 链修正的 Chaos-SVR 时间序列预测模型更优良。

4 结 论

针对办公建筑运行能耗预测问题,本文根据混沌理论构造非线性映射对办公建筑运行能耗时间序列进行相空间重构,提出 Chaos-SVR 的组合预测模型,并利用 Markov 链对组合模型中累积误差进行修正的方法。并将预测模型应用于西安某实际的办公建筑,得出以下结论:

- 1)用 C-C 法求解办公建筑总能耗时间序列嵌入维数和延迟时间,对其进行了相空间重构。同时根据小数据量法求解最大 Lyapunov 指数值为 0.007 7,由此判定办公建筑能耗时间序列存在混沌特性。
- 2)将常用的预测方法比如 Narnet,SVR 与本文所建立方法 Chaos-SVR 预测结果进行比较,本文所建立方法预测误差最小,RMSE 为 6.326 1,MAPE 为 0.287 5,利用 Markov 链对组合模型累积误差修正之后的 RMSE 和 MAPE 分别降为 0.565 5 和 0.044 1。结果表明,Markov 修正之后的预测精度显著提高,为办公建筑的优化运营与节能管理提供充分的决策依据,符合工程实际应用。

参考文献:

- [1] Daniel S, Carlos R, Juan J, et al. Towards the quantification of energy demand and consumption through the adaptive comfort approach in mixed mode office buildings considering climate change[J]. Energy and Building, 2019, 187(0378-7788): 173-185.
- [2] Meng A, Ge J, Yin H, et al. Wind speed forecasting based on wavelet packet decomposition and artificial neural networks trained by crisscross optimization algorithm[J]. Energy Conversion and Management, 2016, 114:75-88.
- [3] Hernandez L, Baladron C, Javier M, et al. Artificial neural networks for short-term load forecasting in microgrids environment[J]. Energy, 2014, 75:252-264.
- [4] Ahmad T, Chen H, Guo Y, et al. A comprehensive overview on the data driven and large scale-based approaches for forecasting of building energy demand: a review[J]. Energy and Buildings, 2018, 165:301-320.
- [5] Ahmad T, Chen H. Short and medium-term forecasting of cooling and heating load demand in building environment with data-mining based approaches[J]. Energy and Buildings, 2018, 460-476.
- [6] Saeed R, Zhang X, Mahdiyar A. A comprehensive review on the application of artificial neural networks in building energy analysis[J]. Neurocomputing, 2019, 340(0925-2312):55-75.
- [7] 应张驰,陈淑萍,卢旭航.基于多源信息的短期负荷混合预测模型应用研究[J].浙江电力,2019,38(9):100-104.
Ying Z C, Chen S P, Lu X H. Application research of short-term load mixed prediction model based on multi-source information[J]. Zhejiang Electric Power, 2019, 38(9):100-104. (in Chinese)

- [8] Ma Z, Ye C, Li H, et al. Applying support vector machines to predict building energy consumption in China[J]. Energy Procedia, 2018, 152(1876-6102):780-786.
- [9] Guo Q, Tian Z, Ding Y, et al. An improved office building cooling load prediction model based on multivariable linear regression[J]. Energy and Buildings, 2015, 107: 445-455.
- [10] Li Q, Meng Q, Cai J, et al. Predicting hourly cooling load in the building: a comparison of support vector machine and different artificial neural networks[J]. Energy Conversion and Management, 2009, 50(1):90-96.
- [11] Zhao J, Liu X. A hybrid method of dynamic cooling and heating load forecasting for office buildings based on artificial intelligence and regression analysis[J]. Energy and Buildings, 2018, 174:293-308.
- [12] Cheng A, Jiang X, Li Y, et al. Multiple sources and multiple measures based traffic flow prediction using the chaos theory and support vector regression method[J]. Physica A: Statistical Mechanics and its Applications, 2017, 466: 422-434.
- [13] Zhou X, Zi X, Liang L, et al. Forecasting performance comparison of two hybrid machine learning models for cooling load of a large-scale commercial building[J]. Journal of Building Engineering, 2019, 21:64-73.
- [14] 董子晗. 基于混沌时间序列的地区电网负荷预测[J]. 电网与清洁能源, 2019(5):38-41.
Dong Z H. Regional power grid load prediction based on chaotic time series[J]. Power Grid and Clean Energy, 2019(5): 38-41.(in Chinese)
- [15] 华琦, 王雷, 陆金桂, 等. 基于混沌理论与神经网络的短期风速滚动预测[J]. 电测与仪表, 2013, 50(7): 16-20.
Hua Q, Wang L, Lu J G, et al. Short-term wind speed rolling prediction based on chaos theory and neural network [J]. Electrical Measurement and Instrumentation, 2013, 50(7):16-20.(in Chinese)
- [16] 行鸿彦, 龚平, 徐伟. 嵌入窗方法确定混沌系统重构参数的仿真研究[J]. 系统仿真学报, 2013, 25(6): 1219-1225.
Xing H Y, Gong P, Xu W. Simulation study of embedding window method to determine reconstruct ion parameters of chaotic system[J]. Journal of System Simulation, 2013, 25(6):1219-1225. (in Chinese)
- [17] Kim H S, Eykholt R, Salas J D. Nonlinear dynamics, delay times and embedding windows [J]. Physica D: Nonlinear Phenomena, 1999, 127: 48-60.
- [18] 孙义, 黄显峰. 基于最大 Lyapunov 指数的混沌预测在洪水实时预报中的应用[J]. 水利水电技术, 2016, 47(1): 102-106.
Sun Y, Huang X F. Application of chaos prediction based on maximum Lyapunov index in real-time flood prediction[J]. Water Conservancy and Hydro-Power Technology, 2016, 47(1):102-106.(in Chinese)
- [19] Torshabi A E, Pella A, Riboldi M, et al. Targeting accuracy in real-time tumor tracking via external surrogates: a comparative study[J]. Technology in Cancer Research & Treatment, 2010, 9(6):551-562.
- [20] 刘义才, 刘斌, 石安伟. 区间化随机时延的网络控制系统建模与控制[J]. 系统仿真学报, 2018, 30(2): 654-663.
Liu Y C, Liu B, Shi A W. Modeling and control of interval stochastic delay network control system[J]. Journal of System Simulation, 2018, 30(2):654-663.(in Chinese)
- [21] 黄银华, 彭建春, 李常春, 等. 马尔科夫理论在中长期负荷预测中的应用[J]. 电力系统及其自动化学报, 2011, 23(5): 131-136.
Huang Y H, Peng J C, Li C C, et al. Application of Markov theory in medium and long-term load forecasting[J]. Journal of Power Systems and Automation, 2011, 23(5):131-136.(in Chinese)
- [22] 陆欣, 沈艳霞, 陈杰, 等. 考虑风力发电随机性的超短期风电功率区间预测研究[J]. 太阳能学报, 2017, 38(5): 1307-1315.
Lu X, Shen Y X, Chen J, et al. Study on the prediction of ultra-short-term wind power power interval considering the randomness of wind power generation[J]. Journal of Solar Energy, 2017, 38(5):1307-1315.(in Chinese)

(编辑 詹燕平)